

# Where to open the most popular Cocktail bar in the town?

## 1. Introduction

### 1.1 Background

In the gastro environment bars of all sorts play a major role. From a simple pub serving beer and maybe some food to exclusive cocktail bars the array and style is unique to every venue. While the venue itself is of great importance to maintaining a solid customers basis, finding new customers can be greatly influenced by the location and neighbourhood. This study focuses on a regional study within Delhi to open a cocktail bar. No food will be served. While the venue wants to capture a high number of high-end customers who will travel to the bar specifically, the goal is to grow their customer base based on location. Essentially making it more accessible and favourable to a broader audience. All business aspects like pricing and products on offer are neglected, focusing solely on the most favourable location to supplement the venue while still making it a standout establishment.

### 1.2 Problem

Data that might contribute for selecting cocktail bar location in Delhi might include bars, restaurants and metro stations in the proximity. As the bar will not serve any food and be a unique venue for one or two drinks in the afternoon/evening following information were taken into consideration:

- Proximity to an array of restaurants
- Reduced competition in the region

## 2. Data

Using the Foursquare API the locations of restaurants, bars, and metro-stations were selected. Delhi is made up of 11 district hence 11 clusters of restaurants, public transport and rival bars were created. The clusters center would serve as the most favourable or least favourable location. Based on the data it might be possible to view regions most fruitful to open the cocktail bar at.

**Optimally regions with high densities of restaurants and public transport and low densities of bars would be recorded.** To also view the accessibility for an after work drink public transport point metro stations were clustered.

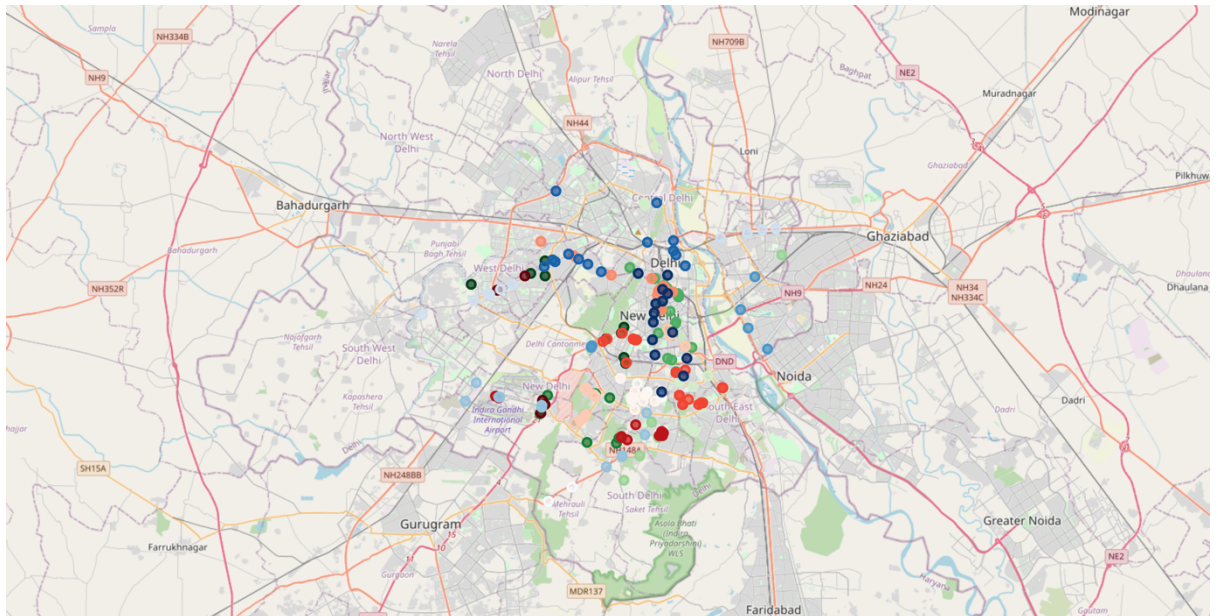
## 3. Methodology

As a database, I used GitHub repository in my study. My master data which has the main components Restaurants, Bars, Metro Stations, Latitude and Longitude information of the city.

	Name	Lat	Lng	Cluster Nr._m
0	Delhi Aerocity Metro Station	28.548572	77.120876	1
1	Dwarka Metro Station	28.614938	77.022775	6
2	New Delhi   नई दिल्ली Metro Station	28.642803	77.223351	10
3	Patel Chowk Metro Station	28.622511	77.213679	10
4	Indira Gandhi International Airport Metro Station	28.555696	77.086223	1

\*Sample output of metro-station data

I used python **folium** library to visualize geographic details of Delhi and I created a map of Delhi with bars, restaurants and metro stations superimposed on top. I used latitude and longitude values to get the visual as below:



I utilized the Foursquare API to explore the restaurants, bars and metro-stations and segment them. I designed the limit as **100 venue** from their given latitude and longitude information. Here is a head of the list Venues name, category, latitude and longitude information from Foursquare API.

#### Restaurants information

	Name	Lat	Lng
0	Thai High	28.522127	77.181831
1	Kainoosh	28.542631	77.156886
2	Olive Bar & Kitchen	28.526103	77.183976
3	The Big Chill Cafe	28.528201	77.217748
4	Maxims	28.552849	77.242065

#### Bar information

	Name	Lat	Lng
0	The Lodhi Hotel	28.591669	77.238131
1	PCO	28.561635	77.155590
2	Lantern's	28.643183	77.177746
3	On The Waterfront	28.592061	77.237951
4	Blue Bar	28.595403	77.170529

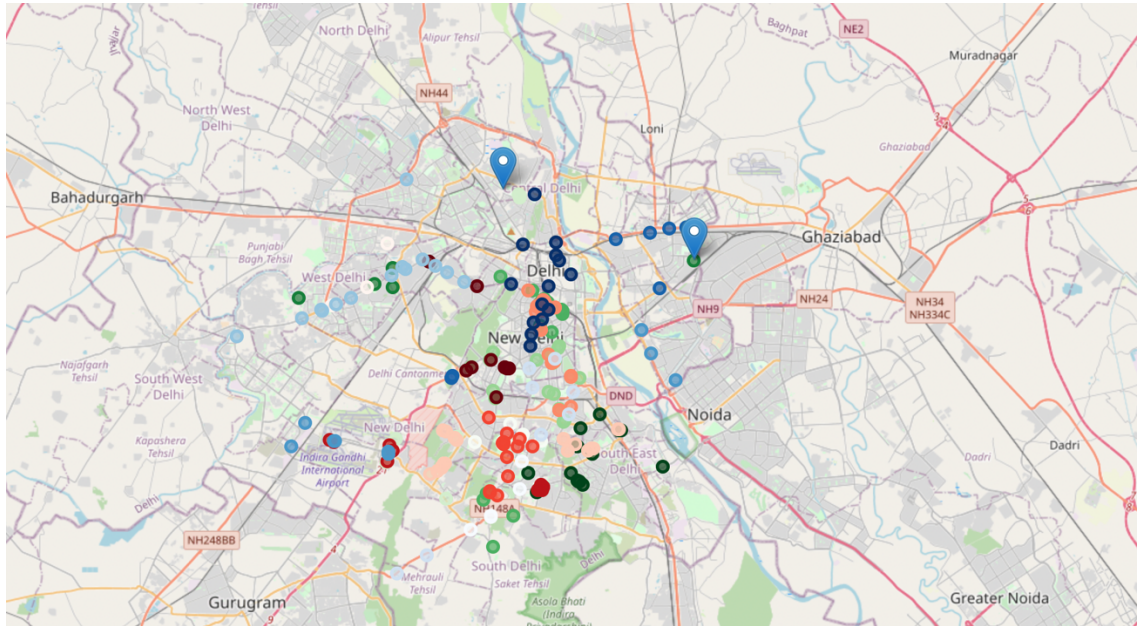
#### Metro-Station information

	Name	Lat	Lng
0	Delhi Aerocity Metro Station	28.548572	77.120876
1	Dwarka Metro Station	28.614938	77.022775
2	New Delhi   नई दिल्ली Metro Station	28.642803	77.223351
3	Patel Chowk Metro Station	28.622511	77.213679
4	Indira Gandhi International Airport Metro Station	28.555696	77.086223

We have some common venue categories. In this reason I used unsupervised learning **K-means algorithm** to cluster the bar, restaurants and metro-stations data into 11 clusters as Delhi has 11 districts. K-Means algorithm is one of the most common cluster method of unsupervised learning.

## 4. Results

As you can see that Delhi city has been divided into 11 clusters and Restaurants – Green, Bars – Red and Metro-stations - Blue represented in the Map and when we place marker at each clusters it can be seen that there are two-three locations are there where bar density is low but restaurants density is there also metro-station connectivity is there based on the above details central Delhi is one the suitable location for opening cocktail bar.



## 5. Discussion

As I mentioned before, Delhi is a big city with a high population density of bars, restaurants in narrow area. The total number of measurements and densities of the 11 districts in total can vary. As there is such a complexity, very different approaches can be tried in clustering and classification studies. Moreover, it is obvious that not every classification method can yield the same high quality results for this metropole.

I used the Kmeans algorithm as part of this clustering study. When I tested the Elbow method, I set the k value to 11. However, only top 100 restaurants and bars coordinates were used. For more detailed and accurate guidance, the data set can be expanded and the details of the neighbourhood or street can also be drilled.

I ended the study by visualizing the data and clustering information on the Delhi map.

## 6. Conclusion

As a result, people are turning to big cities to start a business or work. For this reason, people can achieve better outcomes through their access to the platforms where such information is provided.

Not only for investors but also city managers can manage the city more regularly by using similar data analysis types or platforms.